A discussion on stochastic unfolding

Dijkstra, L.; Eijk, C. van der; Molenaar, I.W.; van Schuur, Hendrik; Stokman, F.N.; Verhelst, N.

Published in:
M.D.N. : methoden en data nieuwsbrief van de Sociaal Wetenschappelijke Sectie van de V.V.S.

Publication date:
1980

Citation for published version (APA):
Dijkstra, L., Eijk, C. V. D., Molenaar, I. W., van Schuur, H., Stokman, F. N., & Verhelst, N. (1980). A discussion on stochastic unfolding. *M.D.N. : methoden en data nieuwsbrief van de Sociaal Wetenschappelijke Sectie van de V.V.S.*, *5*, 158-172.

A DISCUSSION ON STOCHASTIC UNFOLDING

L. Dijkstra (T.H. Eindhoven)
C. van der Eijk (Univ. v. Amsterdam)
I.W. Molenaar (R.U. Groningen)
W.H. van Schuur (N.I.A.S. Wassenaar)
F.N. Stokman (R.U. Groningen)
N. Verhelst (R.U. Utrecht)

0. Introduction

The 1979 volume of M.D.N. contains a paper by Van Schuur and Stokman, in which they suggest that pairwise preference data for six political parties can be analysed by the Mokken scale model (Mokken, 1971). The role of an item would be played by a pair of parties, or rather by their midpoint on the underlying J-scale. Mokken compares the number of error responses for a pair of items with its expectation under the hypothesis of independence. Van Schuur and Stokman modify this expectation into one in which intransitive answer patterns for the six parties are excluded.

Van der Eijk and Van der Noort criticize this modification in a subsequent paper, after which Molenaar wrote a rejoinder on the discussion so far. In December 1979 a one-day conference on the problems raised in this discussion was organized by the Department of Social Sciences at the University of Amsterdam. After this a workgroup was formed by the authors of this paper. Due to lack of time Van der Noort wasn't able to participate. The group discussed several problems and possibilities with respect to the proposed models. Rather than submitting several individual contributions, it was decided that a joint paper be written.

Although the authors agree on the relevance of each of the topics discussed in the following pages, they do not necessarily agree on the relative importance of all the propositions put forward.

1. Stochastic transitivity as a model test.

In an interesting masters' thesis, Jansen (1979) develops a Rasch homogeneous unfolding model (RHUM), in which he applies Rasch analysis to midpoints of stimuli like Sixtl (1973) has also done. In this same thesis, Jansen discusses stochastic transitivity as a model test, as was suggested by Coombs (1964) and Bechtel (1968). The models of Coombs and Bechtel make different predictions on the kind of stochastic transitivity the data must display.

In Coombs' approach the stimuli as well as the individual ideal points are characterized by a probability density function. The (within-subject) variance of the ideal point-distributions is small compared with the variability of, say, the means of the stimulus distributions, so that, for all practical purposes, there is an overlap of a particular ideal point distribution with only a few stimulus distributions. (see figure 1)

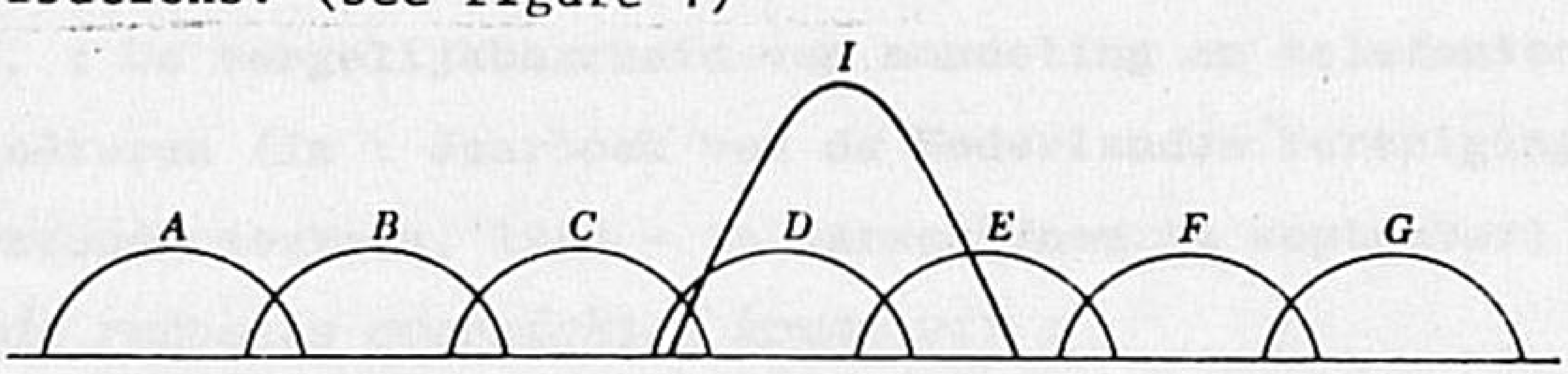

Figure 1. Stochastic unfolding as Coombs views it.
(A,...,G represent stimuli ; I represents an ideal point)

From figure 1, two clear-cut cases follow :

1. $p(A \triangleright B|I)$ depends[+)] only on the degree of overlap between the distributions A and B, but is independent of the momentary value of the ideal point.

+) with $\triangleright$ we mean the empirical relation 'is preferred to', 'is ordered higher than' and the like.

2. p(D ▸ E|I) is dependent on the momentary value of the ideal point, as well as on the momentary values of D and E.

Evidently , as long as the distributions of both stimuli of a pair are situated at the same side of, and without overlap with the distribution of a particular ideal point I (so-called unilateral triples), the pairwise preference probability (of individual I) is governed solely by the stimulus distributions; the variability, and even the location of the ideal point is irrelevant. Stated alternatively: for unilateral pairs the preference model of Coombs can be treated as an application of the law of comparative judgement.

In Bechtel's approach, the stimuli are fixed on the J-scale, and the ideal point is considered as a random variable over subjects and/or over replications (variation within and between subjects). (see figure 2)

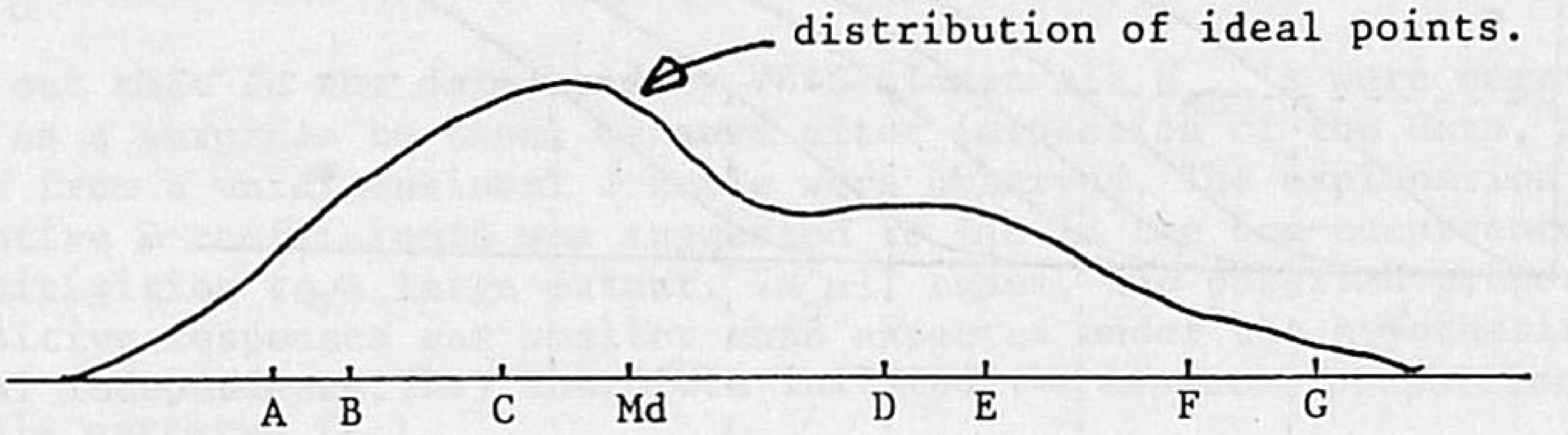

Figure 2. Stochastic unfolding as Bechtel views it.
(A,...,G are (fixed) stimulus values ; Md = Median of ideal point distribution.)

As Greenberg (1965) pointed out already, the order of the pairwise preference probabilities is dependent on the order of the midpoints between pairs of stimuli. Letting AB, AC and BC represent midpoints, then

(1) $p(A \triangleright B) < p(A \triangleright C) \Leftrightarrow AB < AC$

The above equivalence makes sense only in cases where the midpoints between stimuli fall within the range of the aggregate ideal point-distribution, which means that the model is testable only if the variance of the ideal point distribution is of the same order as the stimulus variability. In most practical applications, that will be the case when the ideal point is a random variable over (a heterogeneous group of) individuals.

Now consider unilateral triples. According to Coombs unilateral triples are triples of stimuli whose distribution is situated at the same side of the ideal point distribution I. According to Bechtel they are triples of stimuli whose values lie at the same side of the median of the ideal point distribution. (e.g. stimuli A, B and C in both figure 1 and figure 2.) For a unilateral triple, Coombs and Bechtel predict a different form of stochastic transitivity.
We define

1) Strong stochastic transitivity (SST) as :

$$\text{if } p(A \triangleright B) > .5 \text{ and } p(B \triangleright C) > .5 \text{ then}$$
$$p(A \triangleright C) > \max(p(A \triangleright B), p(B \triangleright C))$$

2) Moderate stochastic transitivity (MST) as :

$$\text{if } p(A \triangleright B) > .5 \text{ and } p(B \triangleright C) > .5 \text{ then}$$
$$p(A \triangleright C) > \min(p(A \triangleright B), p(B \triangleright C))$$

3) Exactly **MST** is defined as MST but not SST.

Coombs model predicts SST for all unilateral triples, as can be understood from the fact that the two outmost distributions have the least mutual overlap, and thus are best discriminated. Bechtels model on the contrary predicts exactly MST for all unilateral triples. This follows from (1).

Jansen points out that his RHUM-model, although derived from different assumptions, makes the same predictions as Bechtels model. These assumptions can best be illustrated by the model Van Schuur and Stokman proposed (1979,p.7 and figure 3 here)

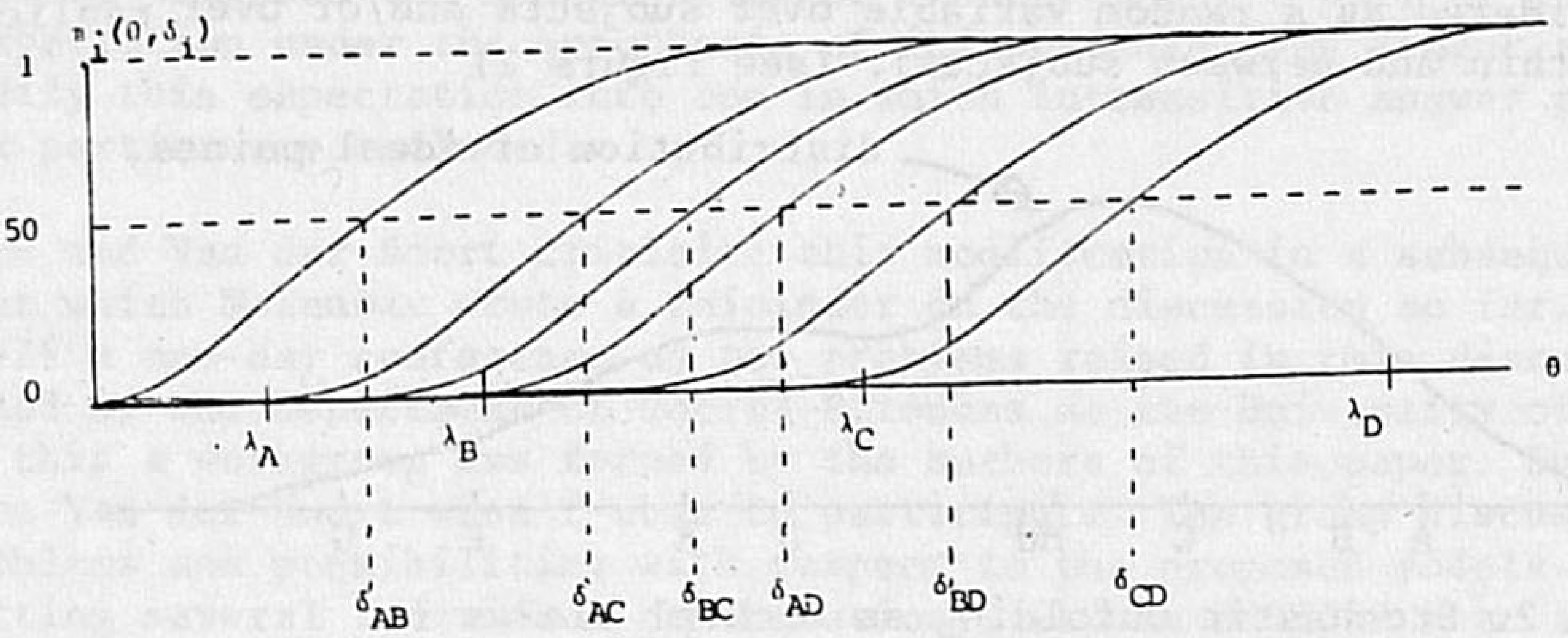

Figure 3. Stochastic unfolding as Van Schuur & Stokman and Jansen view it. (to VS&S, the curves are double monotone functions ; to Jansen they are one-parameter logistic)

When the trace lines for the midpoints are double monotone or even logistic as in the RHUM-model, the trace line for the midpoint AC is placed between those of the midpoints AB and BC. Hence Van Schuur and Stokman as well as Jansen predict exactly MST for unilateral triples. In these two models, the stochasticity is found in the probabilistic response function of the subject to the stimuli. No specific random model of stimulus points and/or ideal points is stated in these models.

For the VS&S data (for details, see VS&S, 1979, p. 22-26) we tried to assess whether or not they conform to the prediction of MST which follows from the VS&S model. To do so we had to construct subgroups of respondents homogeneous in their position on the underlying J-scale. This is necessary because we noticed that unilateral triples can only be defined for a given position of the ideal-point. We constructed subgroups on the basis of current vote intention. For each subgroup we inspected the dominance matrix.

The results can be found in the appendix. It turns out that all unilateral triples (e.g. the triple CDA-VVD-SGP for PvdA voters or the triple CDA-PvdA-PPR for VVD voters) conform to SST rather than to MST. In fact, almost all triples conform to SST. For subjects preferring CPN or SGP however, the dominant I-scale does not conform to the J-scale found for the other subjects (for CPN voters : CPN-PvdA-PPR-VVD-CDA-SGP ; for SGP voters : SGP-CDA-VVD-PvdA-PPR-CPN).

These MST/SST results cast some doubt on the applicability of the models suggested by Bechtel, Sixtl, Jansen and Van Schuur & Stokman to political party preferences.

+) for n stimuli, the dominance matrix is a n by n matrix containing the frequencies with which the column stimulus is preferred to the row stimulus.

About the null model by Van Schuur and Stokman

For the elementary J-scale ABC there are four inadmissible experimental outcomes (from the view of the unfolding model), namely, two outcomes representing intransitive choices, and two inadmissible orderings, ACB and CAB.

The expectation under the hypothesis of statistical independence of the patterns ACB and CAB was found in the Van Schuur and Stokman approach by multiplying p(AB) by p(CB), where p(AB) denotes the probability, in their model, that stimulus A is preferred to stimulus B by a randomly chosen subject.

The coefficient of scalability was defined as

$$H_{ABC} = 1 - \frac{E}{E_0} \tag{2}$$

where E denotes the observed proportion of occurrence of the patterns ACB and CAB, and E_0 denotes the expectation of these patterns, as described above.

It turned out that in the data used by VS&S almost all H_{ABC}'s were negative. This came as a surprise to them, because after inspection of the data, only minor deviations from a unidimensional J-scale were observed. The explanation for these negative H-coefficients was suggested to lie in the non-occurrence of intransitivities to a large extent. In all cases, the observed proportion of intransitive responses was smaller than expected under the hypothesis of statistical independence. They therefore inflated the expected proportion of inadmissible patterns (E_0).

VS & S wanted to make virtue of need, first by deleting those few respondents showing intransitivities, and secondly by inflating E_0 by the proportion of intransitivities as expected under the null hypothesis of statistical independence. In a formula

$$E_0 = \frac{p(AB)\,p(CB)}{1-p(AB)p(BC)p(CA)-p(BA)p(CB)p(AC)} \tag{3}$$

By using this approach, the coefficient of scalability would also be applicable in cases of complete rank order data, rather than pairwise preferences.

Van der Eijk and Van der Noort and later Molenaar showed that something was wrong with formula (3). At the Amsterdam conference it was pointed out that the proportion of orderings where A is preferred to B is not a good estimator of the theoretical (i.e. in the original model of VS & S) probability p(AB). In fact, for a triple of stimuli A,B and C, the sample space (say Ω) consists of 8 points, 6 points representing a transitive choice, and thus an ordering, and 2 points resulting in an intransitivity (i.e., A▷B, B▷C, C▷A and B▷A, C▷B, A▷C). The choice probabilities in the model of VS & S are defined with respect to Ω. Restricting the sample space to the set of transitive choices (say Ω^*) by a process of repeated sampling as suggested by Molenaar, has the consequence that the proportion of orderings where A is preferred to B now is an estimator of a conditional probability, the condition being the restriction of the sample space from Ω to Ω^*.

These considerations make it possible to obtain an estimate of the probabilities p(AB) under the hypothesis that the data are transitive, i.e. that the subjects performed an internal check on transitivity. Let σ(AB) denote an ordering where A is preferred to B, then

$$(4)\ p(\sigma(AB)|\Omega^*) = \frac{p(\sigma(AB)\cap\Omega^*)}{p(\Omega^*)}$$

$$= \frac{p(AB)p(AC) + p(AB)p(CA)p(CB)}{1-p(AB)p(BC)p(CA)-p(BA)p(CB)p(AC)}$$

One can construct a similar expression for the pairs BC and AC. Upon substituting the lefthand members in these expressions by the corresponding observed proportions, the system of three equalities can be solved, giving estimates of p(AB), p(BC) and p(AC). That is precisely what Molenaar and Lewis did[+). The results of their calculations were rather disappointing, because under certain conditions the solution yields rather trivial or implausible results (Molenaar,1979, p. 122).

Although the method used by Molenaar is not optimal (the estimate of the theoretical p(AB) is dependent on which midpoint is used as the third one), the overall negative H's that resulted don't give hope that a pooling method would give better results. The results they obtained cast doubt on the hypothesis of the internal transitivity check, or more accurately, on the compatibility of that hypothesis with the model VS&S proposed.

3. About the null model by Van der Eijk and Van der Noort.

In their own construction of a null-case Van der Eijk and Van der Noort aim primarily at a model applicable for the analysis of rank-order data instead of pairwise preferences. They therefore suggest that complete rank orders of three stimuli should be taken as the elementary datum, rather than pairwise probabilities. For the triple ABC a subject can have one of the following six values: 'ABC', 'ACB', 'BAC', 'BCA', 'CAB' and 'CBA', of which the values 'ACB' and 'CAB' are inadmissible in the J-scale ABC. VdE&VdN use the same coefficient of scalability (2) as VS&S, they also use the same definition of E, but they differ in their definition of EØ. EØ is found by VdE&VdN in the following way:
The observed dominance matrix of the 3 stimuli is the target of their work. This dominance matrix can be thought of as the result of a number of different distributions of subjects over the six response patterns.

By means of linear programming all possible distributions are found. For each distribution, the frequency of occurrence of the patterns ACB and CAB are given. VdE&VdN define EØ as the average frequency of the patterns ACB and CAB over all possible distributions. On pages 36-38 of their original paper they give a numerical example.

Although the VdE&VdN null-model is designed for rank-order data, it is applicable as well for pairwise preferences without intransitivities, as in the data used for analysis by VS&S. However, knowing that pairwise preference questions have been used in collecting such data, the model would imply that a response to a pair of stimuli is deduced from a complete rankorder in the subject's mind. The VdE&VdN model seems less suitable for analysing pairwise preferences with many intransitivities.

+) In their computer program they used the correct formula t=1-y+xy+yz-xz instead of the misprint t=1-y+xy+yz+xz. (Molenaar,p.122)

Some points of criticism to this approach (Molenaar, p. 123) can be set aside as misperceptions once it is noticed that:

i. This approach does not need the assumption that pairwise preferences are given independently;
ii. As complete rank orders of three stimuli are taken as the elementary datum, this approach should not be compared to the calculation of expected frequencies in the 2x2 table formed by 2 stimuli-pairs.

Another criticism, namely that in the construction of all possible distributions only integer-valued distributions over the six patterns are taken into account, is still open. An alternative would be to consider a continuous five-dimensional distribution within the constraints set by the proportions in the dominance matrix. VdE&VdN expect - without having had time to prove so - that this alternative yields the same result as the 'integer-valued' procedure.

4. General problems of model formulation

In the pairwise preference models discussed so far, a subject was imagined either first to produce a perfect rank order in his head and then answering each pair accordingly (Van der Eijk & Van der Noort), or first get tentative answers to each pair and then checking his answers for transitivity (VS&S). Now that we come to think more closely about this, all models involving the choice of a rank order or involving consistency checks may be too far from the mental processes of an actual subject. Still, we see no way around such a description of the preferential choice process.

Problems really get insurmountable when we turn from a three stimulus J-scale to a four stimulus J-scale. Imagine that four stimuli A B C D are to be ordered with the order-obtaining strategy described above.
In terms of the original model of VS&S the sampling space Ω now consists of $2^6 = 64$ points (each point corresponding to 6 independent binary decisions), whereas the sampling space of transitive responses consists of $4! = 24$ points. The 40 intransitive patterns which are possible in this experiment are *not* simply the union of the intransitive patterns in the 4 three-stimulus experiments that can be constructed with the four stimuli.
For example, the 2 three stimulus orderings A B C and B A D, both admissible in a three stimulus experiment do not lead to an ordering in a four stimulus experiment. When one sticks to the hypothesis of repeated sampling with more than 3 stimuli, then the definition of Ω^*, and consequently the denominator in formula (3), must be adapted (and gets very complicated). But apart from the resulting numerical problems , there are other difficulties too.
In the case of n stimuli the number of admissible patterns is n! and the total number of patterns is $2^{\frac{1}{2}n(n-1)}$. The proportions of admissible patterns, being 0.75 for 3 stimuli goes down to 0.02 for 6 stimuli and decreases further at an exponential rate, being only 10^{-7} for n = 10. That implies that a full application of the repeated sampling procedure would take enormous response times, a false prediction as one knows. Another possibility, e.g. in the case of pairwise presentation of the stimuli, would be a consistency check conditional on the responses already given, but in that case the whole experiment is a path-dependent (dependable on the order of presentation), subject-controlled (dependable on the responses actually given) process, and such processes are mathematically very untractable.
Finally, it must be said that the 'repeated sampling hypothesis' is rather gratuitous, since no other evidence than a low proportion of intransitivities, is brought up as an argument. Luce (1959, p. 70-72) already pointed out that several rank ordering strategies are possible, all leading to a different pairwise preference probability estimate. And the problem is especially sharp to VS&S, because their model requires stochastic independence between

all pairwise choices, whereas a rank ordering strategy needs in any case the functional dependence of transitivity. The question if, and how functional dependence and stochastic independence can go together is not easily solved.

The VdE&VdN "explanation" for transitivity in pairwise preferences is also a problematic one when the number of stimuli increases. In a 3-stimuli case it is easily conceivable that pairwise preferences are deduced from a complete rankorder in a subject's mind. In a case with many stimuli (say K) such a hypothesis would require the calculation of expected frequencies (by means of linear programming) of all K! response-patterns. This is not possible because of algorithmic problems. Hence the suggestion (VdE&VdN, 1979, p. 36) to use this approach only for triples, and combining the results in the same way as VS&S do. We see that algorithmic problems lead to an inconsistency in this approach: transitivity is explained by the premise of a complete rank-order in a subject's mind, but this premise is not taken seriously as far as the complete K-tuple is concerned, but only for triples.

5. Data problems

It is highly questionable that we can ignore the problem of functional dependence when we arrive at many more transitive rank orderings on the basis of pairwise preferences than would be expected on the basis of statistical independence. Norpoth (1979), in a similar context as VS&S (pairwise preferences for five West German political parties), remarks that '... surprisingly few individuals were caught in intransitive rankings - no more than 3% of those making all ten comparisons...'. This seems to be especially true when the stimuli are clearly recognizable and distinct objects, such as political parties. We expect these dependency problems to arise less often with less easily recognizable stimuli, as for instance in the case of psychophysical experiments.

Returning to the original VS&S data set, life becomes even more difficult if there are ties and missing data, so when subjects have not produced a perfect rank order. In the actual data, a few persons have tied all fifteen pairs of political parties (a polite way of telling the interviewer to go to hell), a few more have tied all pairs not involving their most preferred party and still more use ties occasionally (mostly among pairs not involving their first choice). As yet we have no satisfactory and simple treatment for such ties.

Moreover, the data contain some indication that unidimensional unfolding may be unsatisfactory regardless of the models we propose. Although many voters agree on a latent ordering CPN-PPR-PvdA-CDA-VVD-SGP, especially among CPN and SGP voters we find systematic deviations from this J-scale. Preferably, such phenomena should be studied when they stand out as residuals, as idiosyncratic individuals or stimuli that fail to obey a model largely confirmed by the others.

6. On the usefulness of characteristic monotony.

When in a dominance matrix the rows and columns are ordered according to the order of the J-scale, then the entries above the main diagonal decrease (weakly) monotonically from left to right and from top to bottom. This phenomenon we call 'characteristic monotony'.

Van Schuur and Stokman propose the use of the property of characteristic monotony (CM) in a dominance matrix in the selection of possible J-scales. They do this in a fairly liberal sense, but they are also convinced that CM is a necessary condition for a stochastic J-scale. Therefore the usefulness of CM in stochastic unfolding is examined in this section.

In the first part the argument for the use of CM is reconstructed. In the second part the conditions are sketched under which CM is disturbed. In the third part some conclusions are drawn with respect to the limitations of CM and in the fourth part a numerical illustration is presented.

6.1 Characteristic monotony.

Given n stimuli, a dominance matrix based on all $\frac{1}{2}n(n-1)$ pairwise preferences has the property of weak monotony if all comparisons fulfill the restriction of transitivity and if the columns and rows are ordered properly. Since ranking n stimuli is identical to making $\frac{1}{2}n(n-1)$ paired preferences under transitivity, a preference pattern can be written as a dominance matrix with the property of weak monotony (this does not imply any specific psychological theory of choice; it is merely a matter of notation).

More specifically : a preference pattern that can be reproduced from a given qualitative J-scale, can be transformed into a dominance matrix with the property of weak monotony by ordering the columns from left to right and the rows from top to bottom corresponding to the left-right order of the stimuli on the J-scale. (Notice that weak monotony in this sense is connected with nothing more than the idea of a <u>qualitative</u> J-scale.)

The sum of a set of dominance matrices like this, of course, is a matrix that has at least the property of weak monotony. The matrix approaches strict monotony dependent on the occurrence of <u>different</u> patterns (their frequency being irrelevant) that can be reproduced from the same qualitative J-scale.

A sum matrix like this, based on a lot of preference patterns that have a qualitative J-scale in common has another important property : the sums of the columns from left to right show a monotone increase to a certain point in the centre and after that a monotone decrease. The rowsums show the opposite figure : from top and bottom they decrease monotonically towards the centre.

If, in reverse, a set of preference patterns has un unknown qualitative J-scale in common, then its identity can easily be recovered by computing a dominance matrix. Rearrangement of rows and columns, resulting in a matrix with the property of (weak) monotony, reveals the stimulus order on the J-scale. Rearrangement is facilitated by the property mentioned above concerning the sums of columns in such a dominance matrix.

6.2 Disturbance of CM.

If in a dominance matrix the rows and columns are ordered corresponding to the stimulus order of a given qualitative J-scale, then the dominance matrix resulting from a preference pattern that <u>cannot be reproduced</u> from this given J-scale, <u>does not have</u> the property of weak monotony.
Furthermore, two mutual mirror image patterns (e.g. ABCD and DCBA) with equal frequencies i result in a summated dominance matrix having i in each cell, irrespective of the ordering of rows and columns.

So, taken seperately, patterns which are not derived from the underlying J-scale disturb (weak) CM of a dominance matrix, but if they are mirror-images and equally frequent, they do not affect CM.

Consequently , a CM matrix will generally result from a subset S of preference patterns that can be reproduced from an identifiable J-scale. The relative share of S in the total set of cases , however, can still vary enormously. +)

Two characteristically different situations can be distinguished in this respect The first situation is the subset S relatively dominating the remaining cases; the reproducibility of the model now mainly depends on the occurrence of metric intransitivity. The CM matrix however, does not provide any information on this point as was noticed before. Since triples do not suffer from this inconvenience, the coefficient H as adapted by van Schuur and Stokman will presumably attain acceptable values.

The second situation places subset S in a relatively minority position, the remaining cases consisting of pairs of mutual mirror images ++). The reproducibility of the model is small. The H-values even tend to become negative.
(By 'the model' we mean : the qualitative J-scale that is indicated by the CM matrix.)

One can easily simulate situation 2, in which the CM matrix identifies a J-scale without scalability, even with negative H-values. Moreover, situation 2, can take a form in which the majority set of remaining cases not only consists of random pairs of mutual mirror images, but contains a J-scale that is hidden by the CM matrix. (see the illustration in 6.4)

+) it is even possible to construct a nontrivial CM-matrix, resulting exclusively from I-scales not compatible with the underlying J-scale. e.g. the I-scales BDECA and CEDBA are not compatible with ABCDE as a J-scale, but their summated dominance matrix is CM. (see also subsection 6.4)

++) The equal occurrence of orderings and their mirror images is not the only way in which all cells in the dominance matrix are increased by a constant. For example the I-scales ACBD, BDCA, CADB and DBAC give a summated dominance matrix containing 2 in all offdiagonal cells.

6.3 Some conclusions.

Van Schuur and Stokman classify their scaling procedures according to two criteria : the stimulus order on the continuum (fixed or free), and the admissibility of selecting out 'bad' stimuli. The last point is of minor importance with respect to the usefulness of the CM matrix.

With a fixed order it is difficult to see why a CM matrix should play such an important role as even to decide a priori on not computing H-values, which is the heart of stochastic unfolding. Only under situation 1 (subset S in the majority position) is this decision justifiable, but at the same time somewhat superfluous, because in that case high H-values and a smooth CM-matrix are two sides of one coin. However in reality we do not know under which situation we are working. On the contrary, the analysis will generally be done to reveal the kind of situation.

With a fixed order under situation 2, we are misled by the the CM matrix, as soon as fixed order and CM order are different, and the fixed order is(partly) scalable (see illustration ; subsection 6.4)

Without a fixed order the situation is certainly not less complicated, because the CM matrix's 'verdict' is now no longer regulated by some theoretical notion.

These objections presumably become even more serious when leaving out stimuli is admissible. Therefore, there seems to be good reason to put the main weight where it belongs in a stochastic approach : in the error cell. The realisation of this suggestion is indicated in the next subsection.

6.4 A numerical illustration

The preceding considerations will now be illustrated by an example (see Table 1). This table reads as follows: the first preference pattern is BDCAE, B being preferred first, etc.

Table 1
preference rankings (artificial example)
n = 140

	A	B	C	D	E	freq.
1	4	1	3	2	5	20
2	2	5	3	4	1	20
3	3	4	2	1	5	20
4	1	5	2	4	3	20
5	5	1	4	2	3	20
6	3	2	4	5	1	20
7	3	1	5	4	2	1
8	1	2	4	5	3	2
9	4	3	5	1	2	3
10	4	2	5	3	1	1
11	2	1	4	5	3	2
12	1	3	2	5	4	3
13	1	2	3	5	4	1
14	3	1	4	5	2	2
15	3	2	5	4	1	3
16	4	3	5	2	1	2

We take 4 I-scales in relatively large frequencies from J-scale BDCAE, including the mirror images BDCAE (1) and EACDB (2). The two remaining I-scales DCABE (3) and ACEDB (4) are 'neutralized' (from the point of view of monotony) by EBACD (6) and BDECA (5) respectively.
Now we admix in much smaller frequencies nearly all the I-scales that can be reproduced from another J-scale: DEBAC (patterns 7 to 16) (Except the mirror image CABED; this omission, however, is not essential at all for the argument).
The next step consists of computing the dominance matrix and re-arranging it into characteristic monotony (see Table 2). Now H-values can be computed, guided by this CM-matrix.

Table 2

Dominance matrix, rearranged to CM based on the data in table 1

	D	E	B	A	C
D	-	77	75	74	70
E	63	-	71	68	64
B	65	69	-	66	63
A	66	72	74	-	60
C	70	76	77	80	-

We do, however, not just regard the triples that are relevant in this respect. On the contrary, all the possible triples are investigated (see Table 3). J-scale DEBAC, indicated by the CM matrix, fails, as can be seen from Table 3 (consider the x marked cells). On the other hand, looking for error cells containing smaller frequencies than expected under the null hypothesis, the eye is stricken by nearly all the triples compatible with J-scale BDCAE, introduced in large frequencies earlier (consider the + marked cells in Table 3). This conclusion holds regardless of whether the VS&S definition of expected frequencies, or the VdE&VdN one. (N.B.: It can easily be noticed in table 3 that in some cases the expected frequencies according to VS&S and according to VdE&VdN differ only slightly . At this moment it is not evident under which circumstances these approaches lead to different results. This remains a point for further study, and should be taken into account in a more definite evaluation of both approaches).

These results lead to the conclusion (following an idea of Dijkstra (1978)), that the function of the CM matrix in stochastic unfolding, as proposed by Van Schuur and Stokman, probably can be performed better by starting from acceptable triples to be selected out of all possible triples (see Table 3). From this starting point on, building more complex scales does not seem to be difficult in principle. For example, all acceptable triples in Table 3 (excluding ABE and BCE) are compatible with J-scale BDCAE. With the foreknowledge we have, this is not surprising. Error cell frequencies associated with J-scale BDCAE that have not been considered thus far, and their expected frequencies, look like this (calculated according to VS&S, 1979, p. 13):

Table 3

Triples, error-cell frequencies (E), and expected error-cell-frequencies (E_0). Expected error-cell frequencies: upper number: expectation calculated according to Van Schuur and Stokman i.e. corrected for transitivities). Middle number: expectation based directly on marginal totals in table 2 (no correction for transitivities). Lower number: expectation calculated according to Van der Eijk and Van der Noort.

	I		II		III	
triples	E	E_0	E	E_0	E	E_0
ABC	40	42.1 31.7 x 40	63	39.4 29.7 43	37	58.4 44.0 + 57
ABD	46	46.4 34.9 44.3	60	40.8 30.6 x 43.3	34	52.8 39.6 + 52.3
ABE	52	50.7 38.1 + 50	40	43.3 32.5 x 44	48	45.9 34.5 46
ACD	60	37.6 28.3 x 40	10	53.2 40.0 + 50	70	49.2 37.0 50
ACE	20	41.0 30.9 x + 40	56	57.7 43.4 56	64	41.3 31.1 44
ADE	26	45.2 33.9 + 45	74	54.1 40.7 53	40	40.7 30.6 x 42
BCD	60	38.9 29.3 x 42	50	51.2 38.5 49	30	49.9 37.5 + 49
BCE	40	41.3 31.1 x 42	56	55.6 41.8 + 55	44	43.1 32.5 43
BDE	45	42.6 32.0 44	35	54.9 41.3 + 54	60	42.5 32.0 x 42
CDE	30	50.5 30.0 + 49	70	51.2 38.5 49	40	38.3 28.8 x 42

x : corresponding J-scale compatible with J-scale DEBAC

+ : corresponding J-scale compatible with J-scale BDCAE

note : see next page

note concerning table 3:

The notation in table 3 reads as follows : out of the 3 stimuli in alphabetic order (say ABC) in column I the first stimulus (in this case A) is the middle one on a J-scale ; so the J-scale is BAC. In column II the second stimulus (in this case B) is the middle one, and in column III the third stimulus is the middle one. So, with the stimuli ACE, column III means J-scale AEC.

The numbers not between parentheses are the frequencies of the incompatible patterns. For example the patterns of 37 cases that cannot be reproduced from J-scale ACB. It can be shown that the frequencies in table 3 correspond to the concept of error cell frequencies as developed by van Schuur and Stokman. For example, ABC I contains the sum frequency of patterns BCA and CBA, neither of these being compatible with J-scale BAC. This stress frequency of J-scale BAC is identical with the error cell frequency in the van Schuur and Stokman approach, since they define the error cell frequency of J-scale BAC as the number of respondents that did not pass midpoint BA but at the same time did pass midpoint AC. Therefore, in the triple ABC, stimulus A is ranked last by these persons, which is expressed in the patterns BCA and CBA.

midpoints			
not passed	passed	f_o	f_e
BD	CA	35	42.9
BD	CE	71	40.7
BD	AE	47	38.6
DC	AE	30	36.0

Consequently BDCAE is rejected as J-scale. The remaining possible J-scales are now BDCA and DCAE, both having error cell frequencies smaller than the corresponding expected frequencies.

The example presented above is intended to warn against putting too much weight on the property of CM in the (overall) dominance matrix in the construction of a J-scale. It is not intended to propose that CM can be totally discarded as a criterion for the evaluation of J-scales. The property of CM follows directly from the notion of a one-dimensional scale where stimuli (or rather, their midpoints) and subjects have a certain (fixed, or momentary) location. This means that according to this formal model, triples to be selected for the formation of a J-scale on the basis of their H-value, should be CM themselves (i.e. in case of the qualitative J-scale ABC: $p(AB) < p(AC) < p(BC)$). In the example presented above not all triples with satisfactory H-values pass this test. It remains a point for further research to establish whether or not the criterion of CM for the evaluation of a J-scale can be relaxed, and, if so, under what conditions, and to which degree.

7. Conclusion.

The problems discussed in this article, and the problems of stochastic unfolding in general are serious ones. Apart from some minor problems, such as how to treat ties (though not yet solved), we discussed three main problems : the problem of stochastic transitivity, the problem of the transitivity check by subjects, and the problem of the dominance matrix.

(1) All models considered, with the exception of Coombs', predict exactly MST for unilateral triples, whereas the data set of van Schuur and Stokman (along with several others) show SST. Since no acceptable explanation is offered for this failure, maybe it must be considered fatal, leading to the rejection of the 'weak' models (van Schuur and Stokman, Bechtel) as well as the 'strong' ones (Sixtl, RHUM) discussed.

(2) The data analysed by Van Schuur and Stokman as well as those presented by Norpoth show clear evidence of an internal 'transitivity check'. Two strategies are presented to cope with this phenomenon :

(a) The approach of Van der Eijk and Van der Noort in a sense makes the transitivity check the heart of the model, in that they consider the pairwise preferences from an implicit ordering of the stimuli, thus excluding intransitivities a priori. It must be asked if their model, notwithstanding the statistical procedure used, is a stochastic model of individual choice.

(b) The approach by Van Schuur and Stokman, together with the correction for transitivity (formula (3)) leads to the very serious dilemma of simultaneously requiring stochastic independence at the level of pairwise preferences, and functional dependence brought about by the transitivity check.

(3) The third kind of problem encountered is an algorithmic one. As shown in section 6, there is a risk of losing important information when one puts too much weight on the CM-property of the dominance matrix. This risk applies to the VS&S approach, as well as to VdE&VdN. The suggestion made in section 6, namely, to build J-scales, starting from acceptable triples (in fact, this is a stochastic extension of Dijkstra's study on unfolding, and resembles very much Mokken's algorithm of scale construction) can be an elegant solution to this problem.

References.

BECHTEL, G.G. Folded and unfolded scaling from preferential paired comparisons. J. Math. Psychol.,1968 (5), 333-357.

COOMBS, C.H. A Theory of Data, N.Y.,Wiley, 1964.

DIJKSTRA, L. Ontvouwing.Over het afbeelden van rangordes van voorkeur in ruimtelijke modellen, Assen, van Gorcum, 1978.

EIJK, C. van der & NOORT, W. van der Some notes concerning stochastic unfolding and the expected frequency of rankorder patterns M.D.N., 1979, vol. 4, nr. 2, 30-38.

GREENBERG, M.G. A Method of Successive Cumulations for the Scaling of Pair-comparisons Preference Judgments Psychometrika, 1965 (38), 441-448.

JANSEN, P.G.W. The Rasch model and attitude measurement. Nijmegen, 1979 (unpublished master's thesis).

LUCE, R.D. Individual Choice Behavior, N.Y., Wiley, 1959.

MOKKEN, R.J. A Theory and Procedure of Scale Analysis, Den Haag, Mouton, 1971.

MOLENAAR, I.W. Two null models and their flaws. M.D.N., 1979, vol. 4, nr. 3, 120-125.

NORPOTH, H. The Parties come to order. Dimensions of Preferential Choice in the West-German Electorate , 1961-1976 The American Political Science Review, 1979 (73), 724-736.

SCHUUR, W.H. van & STOKMAN, F.N. A one-dimensional stochastic unfolding model with an application to party preferences in the Netherlands. M.D.N., 1979, vol. 4 nr. 4, 3-29 .

SIXTL, F., Probabilistic Unfolding. Psychometrika, 1973 (38), 235-248.

APPENDIX. Dominance matrices for respondents with similar vote intention.

Vote intention was measured by the response to the question : '*If elections for the "Tweede Kamer" would be held now, would you vote at all ? If yes, what party would you vote for ?*'

Table A.1

Dominance matrix of respondents who express a vote intention for CDA. (N=178)

	CDA	VVD	PvdA	SGP	PPR	CPN
CDA	-	12	11	4	5	0
VVD	161	-	80	48	37	10
PvdA	164	83	-	52	28	7
SGP	170	101	104	-	61	13
PPR	172	119	118	76	-	10
CPN	178	144	149	112	114	-

Reading example : 161 out of the 178 respondents who express a vote intention for CDA, prefer CDA to VVD, whereas 12 respondents prefer VVD to CDA.

Apparently 5 respondents (178 - (161 + 12)) did not choose between both parties (i.e. they produced a tie). As the number of ties differs substantially for different pairs of parties in this dominance matrix, we should correct for this before testing on stochastic transitivity. This correction can be accomplished in two ways :

(α) Calculate the proportion for those respondents only who pronounced a choice.

eg.: CDA-VVD : $\frac{161}{161 + 12} = .93$

(b) Allocate the respondents who did not make a choice equally over both alternatives

e.g : CDA-VVD : $\frac{161 + \frac{1}{2}(178 - 161 - 12)}{178} = .92$

In the original S & S paper the qualitative J-scale was found to be : CPN-PPR-PvdA-CDA-VVD-SGP. For respondents who express a vote intention for CDA, the only unilateral triple among the possible triples is PvdA-PPR-CPN. We therefore must consider the order of the probabilities p(PvdA ▸ PPR), p(PvdA ▸ CPN) and p(PPR ▸ CPN)

The 'exactly MST' prediction is p(PvdA ▸ PPR) ≤ p(PvdA ▸ CPN) ≤ p(PPR ▸ CPN)
The SST prediction is :
p(PvdA ▸ CPN) ≥ p(PvdA ▸ PPR) and p(PvdA ▸ CPN) ≥ p(PPR ▸ CPN

As far as the frequencies are concerned, the data confirm the SST prediction (see table A.1). Using the proportion with correction for ties gives the same result, independent of the type of correction (a or b) applied.

ad a.: $\hat{p}$(PvdA ▸ CPN)=.96 ; $\hat{p}$(PvdA ▸ PPR)=.81 ; $\hat{p}$(PPR ▸ CPN)=.92
ad b.: $\hat{p}$(PvdA ▸ CPN)=.90 ; $\hat{p}$(PvdA ▸ PPR)=.75 ; $\hat{p}$(PPR ▸ CPN)=.79

The property of SST for unilateral triples holds for all respondents with identical vote intentions for each of the six parties used. For CPN- and SGP-voters, the exception must be made, however, that their dominant qualitative J-scale on the basis of the dominance matrix does not conform to the overall J-scale. Hence the unilateral triples with parties in different order (as CDA-VVD instead of VVD-CDA as in the overall J-scale) do not conform to the SST-prediction. As the results do not differ when using the frequencies, or one of the two formentioned proportions-corrected-for-ties, we will only present the results with the frequencies.

Table A.2

Dominance matrix of respondents who express a vote intention for PvdA. (N=157)

	PvdA	PPR	CDA	VVD	CPN	SGP
PvdA	-	3	3	2	1	0
PPR	151	-	49	29	13	6
CDA	154	92	-	36	30	13
VVD	154	108	100	-	54	47
CPN	155	120	100	69	-	50
SGP	152	117	111	67	54	-

Table A.3

Dominance matrix of respondents who express a vote intention for VVD. (N=83)

	VVD	CDA	PvdA	PPR	SGP	CPN
VVD	-	3	1	0	1	0
CDA	80	-	12	6	6	2
PvdA	82	62	-	23	28	2
PPR	83	66	39	-	31	5
SGP	81	67	38	31	-	5
CPN	82	70	60	49	51	-

Table A.4

Dominance matrix of respondents who express a vote intention for PPR.(N=16)

	PPR	PvdA	CDA	CPN	VVD	SGP
PPR	–	0	0	0	1	0
PvdA	15	–	2	2	1	0
CDA	16	13	–	5	3	2
CPN	14	12	7	–	4	3
VVD	14	14	9	9	–	4
SGP	14	15	10	9	7	–

Table A.5

Dominance matrix of respondents who express a vote intention for SGP.(N=10)

	SGP	CDA	VVD	PvdA	PPR	CPN
SGP	–	0	1	0	0	0
CDA	10	–	3	2	1	0
VVD	9	6	–	0	1	0
PvdA	10	7	3	–	2	0
PPR	10	8	6	2	–	0
CPN	10	9	5	4	4	–

Table A.6

Dominance matrix of respondents who express a vote intention for CPN.(N=8)

	CPN	PvdA	PPR	VVD	CDA	SGP
CPN	–	1	0	0	0	0
PvdA	7	–	1	1	0	0
PPR	7	6	–	2	1	0
VVD	8	6	3	–	2	1
CDA	8	7	3	2	–	0
SGP	8	7	4	3	3	–